CMPSCI 687: Reinforcement Learning
Fall 2026 Class Notes

Professor Philip S. Thomas
University of Massachusetts
pthomas@cs.umass.edu

Table of Contents

1 Introduction

This document contains notes related to what we cover in class. This is intended as a replacement for posting student notes each time the course is offered (see, for example, the hand-written notes from the 2017 offering here). This is not meant to be a standalone document like a textbook.

1.1 Notation

When possible, sets will be denoted by calligraphic capital letters (e.g., $\mathcal{X}$), elements of sets by lowercase letters (e.g., $x \in \mathcal{X}$), random variables by capital letters (e.g., X), and functions by lowercase letters (e.g., f). This will not always be possible, so keep an eye out for exceptions.

We write $f : \mathcal{X} \to \mathcal{Y}$ to denote that f is a function with domain $\mathcal{X}$ and range $\mathcal{Y}$. That is, it takes as input an element of the set $\mathcal{X}$ and produces as output an element of $\mathcal{Y}$. We write $|\mathcal{X}|$ to denote the cardinality of the set $\mathcal{X}$—the number of elements in $\mathcal{X}$, and $|x|$ to denote the absolute value of x (thus the meaning of $|\cdot|$ depends on context).

We typically use capital letters for matrices (e.g., A) and lowercase letters for vectors (e.g., b). We write $A^\intercal$ to denote the transpose of A. Vectors are assumed to be column vectors. Unless otherwise specified, $\|b\|$ denotes the l^2-norm (Euclidean norm) of the vector v.

We write $\mathbb{N}_1$ to denote the natural numbers starting from 1 (i.e., *not* including zero), and $\mathbb{N}_0$ to denote the natural numbers including zero.

We write $\coloneqq$ to denote *is defined to be*. In lecture we may write $\triangleq$ rather than $\coloneqq$ since the triangle is easier to see when reading my (sometimes sloppy) handwriting.

If $f : \mathcal{X} \times \mathcal{Y} \to \mathcal{Z}$ for any sets $\mathcal{X}$, $\mathcal{Y}$, and $\mathcal{Z}$, then we write $f(\cdot, y)$ to denote a function, $g : \mathcal{X} \to \mathcal{Z}$, such that $g(x) = f(x, y)$ for all $x \in \mathcal{X}$.

We denote sets using brackets, e.g., $\{1, 2, 3\}$, and sequences and tuples using parentheses, e.g., $(x_1, x_2, \dots)$.

We write $\operatorname{supp}(X)$ to denote the support of a random variable X. For discrete random variables, $\operatorname{supp}(X) = \{x : \Pr(X = x) > 0\}$.

The notation that we use is *not* the same as that of the book or other sources (papers and books often use different notations, and there is no agreed-upon standard). Our notation is roughly a mix between the notations of the first and second editions of Sutton and Barto's book.

End of Lecture 1, September 9, 2026

References